

A METHOD OF MANAGING RADIO RESOURCES IN AN INTERACTIVE
TELECOMMUNICATION NETWORK

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application is based on French Patent Application No. 00 13 212 filed October 16, 2000, the disclosure of which is hereby incorporated by reference thereto in its entirety, and the priority of which is hereby claimed under 35 U.S.C. §119.

BACKGROUND OF THE INVENTION

Field of the invention

10 The present invention relates to radiocommunications, for example to interactive telecommunication networks using or integrating satellites, and more particularly to managing the availability of resources and services for the various users of an interactive telecommunication network.

15 The present invention provides a method of managing radio resources, in particular the availability of radio resources, in an interactive telecommunication network, as well as a network and a terminal for implementing the radio resource management method.

Description of the prior art

20 A major question to be addressed by the designers of radiocommunication networks or systems, such as a satellite telecommunication system in particular, relates to the availability of resources and services, which contributes to the quality of service provided to the end user. Accordingly, several aspects have to be considered with a view to predicting propagation parameters, necessary in particular in designing Earth-space systems.

25 ITU-R Recommendation P.618-5 lists the effects to be taken into consideration and the model to be used to take them into account when designing Earth-space systems. A fixed margin is usually established in the clear sky links budget in the worst case scenario of use of the system, so as to anticipate and compensate degradation such as attenuation due to rain or a loss of free space, in accordance with the target availability of the system and the climatic area.

30 In a service area with a diameter of several hundred kilometers, propagation conditions can vary greatly and evolve differently from one place to another within said area covered by the system. A fixed margin implies that the capacity of channels is insufficiently used for a rather high percentage of the time.

35 However, there are situations in which the philosophy of designing a fixed margin into the links budget is highly unsatisfactory, and one characteristic example of

09975960 "101501
this is provided by Ku or Ka band interactive satellite communication systems.

Given the variability of propagation conditions from one place to another, leading to a kind of statistical multiplexing of attenuation, one effective approach is to use a margin shared between active terminals within the service area, with a view to
5 reducing the margin established to counter fading. That margin is generally decided on by assuming a small number of terminals subject to rainy weather within the service area.

However, the above solution does not prevent terminals subject to strong attenuation from affecting the quality of service for terminals connected to the same resource under clear sky conditions, or serious and unbalanced degradation of service
10 when the number of terminals subject to rainy weather exceeds the number used to decide on the margin, which may have been the result of long-term statistical analyses.

The present invention aims in particular to alleviate the drawbacks previously cited through assistance with managing the consumption of the shared global margin with optimum conformance to set points for quality of service and availability of
15 resources for each terminal.

SUMMARY OF THE INVENTION

To this end, the present invention provides a method of managing radio resources in an interactive telecommunication network which includes a plurality of terminals severally sharing the same available radio resource and is preferably of the
20 type including at least one satellite, in which method communication services and resources allocated by the network to a given connected terminal t_i for uplink and/or downlink transmission are managed as a function of the value for the terminal t_i of a product $\alpha^{(i)}$ of the type:

$$\alpha^{(i)} = \text{bandwidth } r_i \times \text{power } p_i,$$

25 with i varying from 1 to N , where N is the total number of terminals in the service area concerned.

The bandwidth term r_i preferably corresponds to the cumulative equivalent bandwidth of the connections of the terminal t_i , estimated at the time of accepting the call or connection concerned and the power term p_i corresponds to the average
30 consumption of the terminal t_i , the value of p_i being determined periodically.

In one embodiment of the invention the allocation of communication services and resources to a connected terminal t_i is a function of the result of comparing the calculated value of the product $\alpha^{(i)}$, subject to a corresponding threshold value $\alpha_{\text{lim}}^{(i)}$ in the form of a maximum bandwidth $r_i \times$ power p_i product, with the quantity of radio
35 resources reserved for accepting the connection, augmented by a supplementary margin

for achieving the availability of service required or fixed for the terminal t_i .

Thus the equivalent bandwidth r_i allocated to the terminal t_i is reduced by a factor $\alpha^{(i)} / \alpha_{\text{lim}}^{(i)}$ if the product $\alpha^{(i)}$ becomes greater than the product $\alpha_{\text{lim}}^{(i)}$ (division of r_i by the latter factor) and, for a terminal t_i having a plurality of connections with different classes of service, the equivalent bandwidth reduction is shared between the various connections at random or in accordance with a predetermined hierarchical order.

Thus priority can be given to reducing the bandwidth(s) of the connection(s) using the most radio resources or those with reduced quality of service, for example.

In the event of serious reduction of bandwidth, the reduced bandwidth may no longer justify maintaining the connection concerned.

Thus the connection can be cut off or cleared down if its equivalent bandwidth falls below a lower threshold value r_{min} corresponding to the minimum binary bit rate allocated to a terminal t_i for the connection concerned, for example.

After the equivalent bandwidth r_i of a terminal t_i has been reduced beforehand, the equivalent bandwidth r_i is progressively returned to its normal value before reduction if the product $\alpha^{(i)}$ becomes less than the product $\alpha_{\text{lim}}^{(i)}$ again.

To take into account also the overall situation in terms of radio resource capacity, a second phase of verification and adaptation, for a given radio resource, such as a carrier, shared by a group G_i of several terminals t_i , the communication services and resources allocated by the network to the terminals t_i of the group G_i are advantageously managed globally as a function of a parameter α^{T_i} defined by the equation:

$$\alpha^{T_i} = \sum_{G(i)} r_i \times p_i.$$

To be more precise, in this additional phase, the equivalent bandwidth r_i of all the terminals t_i of the group G_i is reduced uniformly or in a differentiated manner or in a weighted manner if the parameter α^{T_i} exceeds a threshold value $\alpha_{\text{lim}}^{(T_i)}$ corresponding to the capacity of the common radio resource shared by the terminals t_i of the group G_i .

The equivalent bandwidth r_i of all the terminals t_i of the group G_i is preferably reduced by a factor $\alpha^{T_i} / \alpha_{\text{lim}}^{(T_i)}$ (division of r_i by the latter factor).

To limit the reduction in quality of service to what is strictly necessary, the equivalent bandwidth reduction can be applied in a random or hierarchically predetermined manner to different terminals t_i of the group G_i in succession, the product α^{T_i} is calculated again after each reduction of the equivalent bandwidth r_i for a terminal t_i , and continued application of the equivalent bandwidth reduction to the group G_i halted immediately the following condition is verified:

$$\alpha_{Ti} \leq \alpha_{lim}^{(T_i)}.$$

If the equivalent bandwidths r_i are reduced until they are equal to their respective minimum binary bit rates and the condition $\alpha_{Ti} \leq \alpha_{lim}^{(T_i)}$ is still not verified for all the terminals t_i of the group G_i , all that remains is for the terminals t_i to be disconnected from the network to be chosen at random.

In the most comprehensive embodiment of the method according to the invention, and in a cyclic process, the bandwidths r_i of the terminals t_i connected to the network via the at least one satellite are managed individually at a first stage and the terminals t_i of the groups G_i each associated with a shared radio resource are managed globally or in a grouped manner at a second stage.

The method is preferably applied to a code division multiple access satellite multimedia telecommunication network with automatic matching of the power transmitted from and to each terminal to the propagation conditions.

Implementation of the above method is based on the capacity of the algorithms used to allocate radio resources for transmitting traffic data to apply the reduction. For example, weighted circular permutation algorithms and weighted equitable queuing algorithms are compatible and usable with this kind of method.

It will be noted that the management method according to the invention is based on a plurality of fundamental criteria, with a view to achieving various objectives, namely:

- isolation between the terminals: strong degradation of the propagation conditions of the radio channel of one terminal will not affect the terminals sharing the same radio resource (e.g. the same carrier).
- equity: a lack of capacity will be assumed equally between the terminals.
- guaranteed minimum binary bit rate: even in the eventuality of severe fading, the terminal can maintain a connection at a minimum binary bit rate.

On the basis of the above three points, it is possible to define two type of availability:

- service availability: this is the percentage of the time for which the system or network is accessible with nominal performance. This means that during this time the system must comply with the quality of service requirements.

- system access availability: this is the percentage of the time for which the system or network is accessible, even with degraded performance; the cause of the degraded performance of the link can be degraded propagation, for example caused by

attenuation or scintillation due to heavy rain.

The present invention also provides an interactive satellite radiocommunication network or system providing communication channels and connections to a plurality of fixed or mobile terminals severally sharing the same radio resource made available by the network, wherein communication services and resources allocated to a given terminal t_i for uplink and/or downlink transmission are managed as a function of the value for the terminal t_i of a product $\alpha^{(i)}$ of the type:

$$\alpha^{(i)} = \text{bandwidth } r_i \times \text{power } p_i.$$

The above system preferably uses the management method which is described above and one particular application of which is illustrated hereinafter, and to this end includes means for providing gateways suitable for packet-based multimedia traffic between terminals in different service areas, a central network or a base, radio resource control means providing in particular a connection acceptance control function, a media access control function, and a power control function, and means for managing margins which match equivalent bandwidths continuously or in a stepwise manner during the existence of connections as a function of the corresponding calculated values $\alpha^{(i)}$ and α^{T_i} .

The system advantageously further includes at least one traffic supervisor adapted to redistribute the radio resources allocated to each downlink transmission communication gateway and a dedicated logical signaling interface for each terminal t_i for adapting equivalent bandwidths and transmitting corresponding information to the traffic supervisor means.

Finally, the invention also provides a fixed or mobile telecommunication terminal forming part of a system or network of the type previously cited and adapted to implement the method described herein.

In what follows, the present invention is described in relation to one non-limiting example of its practical application, that application nevertheless constituting a preferred embodiment, integrated into a packet-based satellite communication system based on a generic architecture of the following kind.

A gateway in a service area is responsible for transporting data traffic and/or voice traffic between terminals in the service area and a central network. It accepts packet-based multimedia traffic and employs a plurality of functions to provide a quality of service for transport on the connection. It defines how radio resources are shared between the terminals and how radio resources are controlled in conformance with quality of service requirements.

Three main functions govern the use of radio resources: a connection

acceptance control (CAC) function, a media access control (MAC) function, and a power control function.

5 The connection admission control function controls radio resources and decides whether to accept a connection at connection set-up time. This function assumes a capacity that takes into account a global margin shared by the terminals routed on the same radio resource (carrier or set of carriers), in order to apply fading countermeasures without degrading service.

10 The power control function measures the quality of the link for each terminal and adapts the power demand of each terminal to maintain the required cell error rate on the physical transmission medium.

The media access control function allocates a radio resource on a packet basis. It uses carrier load algorithms, quality of service obligation supervisors, and a return access protocol based on demand assignment multiple access (DAMA), the natures and modes of operation of which are known to the skilled person.

15 On the same principle as before, a fourth function known as the margin manager function has been added to adapt the equivalent bandwidth during the life of the connection in accordance with a strategy described hereinafter.

20 A practical embodiment of the method according to the invention applied to a "gateway to terminal" link is described in detail hereinafter. An embodiment applied to a "terminal to gateway" link is described in less detail.

The power is established on a terminal basis but the equivalent bandwidth is established on a connection basis. The following explanations describe in detail the reduction or the increase in the global equivalent bandwidth for the terminal and the sharing between the connections of a given terminal.

25 I) Gateway to terminal link

The following notation is used hereinafter:

- r_i : equivalent bandwidth of the i^{th} terminal, which is the sum of the equivalent bandwidths of the connections estimated by the CAC function for accepting the call.

30 - $\alpha^{(i)}$: the real bandwidth \times power product for the i^{th} terminal. It is calculated from the average power consumption and is refreshed as soon as a new value of the average power consumption is available.

- $\alpha_{\text{lim}}^{(i)}$: the maximum bandwidth \times power product for the i^{th} terminal. It corresponds to the quantity of radio resources reserved by the CAC function to accept the connection, augmented by the supplementary power margin necessary to achieve service availability for the i^{th} terminal, which can be considered as a virtual individual

35

margin.

- α^T : the global bandwidth \times power product for all of the terminals of the carrier. It is given by the equation $\alpha^T = \sum_{SKT} r_i \cdot p_i$

- α_{lim}^T : the maximum bandwidth \times power product for the carrier. It corresponds

5 to the capacity of the carrier.

- p_i : the average power consumption of the i^{th} terminal. It is calculated periodically from all of the powers provided by the power control function.

- r_{min} : the minimum binary bit rate offered to each terminal.

Capacity reduction

10 The first step individually verifies $\alpha^{(i)}$ against $\alpha_{lim}^{(i)}$. If the current value of $\alpha^{(i)}$ exceeds the limit for a terminal, the associated equivalent bandwidth r_i is reduced by a factor $\alpha^{(i)}/\alpha_{lim}^{(i)}$, either continuously or in discrete steps to simplify the implementation.

15 If the equivalent bandwidth falls below r_{min} , the gateway is no longer capable of maintaining a minimum binary bit rate channel and the terminal is disconnected from the system.

The second step globally verifies the current value of α^T against α_{lim}^T . If the current value of α^T exceeds the limit for the carrier, the equivalent bandwidth for all the terminals of the carrier is reduced by a factor α^T/α_{lim}^T . For terminals below r_{min} , the gateway is no longer capable of maintaining the minimum binary bit rate channel and the terminal is disconnected from the system.

20 This second step is based on the radio hypothesis used to establish the collective margin. This assumes only a small number of terminals exposed to rainy weather, which can be considered as a rainy activity per terminal factor. This hypothesis leads to a collective margin much lower than the individual margin to achieve the same availability. However, the number of "rainy weather" terminals is sometimes greater than the number used to determine the margin. In this case, the collective processing reduces the equivalent bandwidth, even if each "rainy weather" terminal has not used up all of its individual margin.

25 The bandwidth reduction applies to the cumulative equivalent bandwidth of a terminal and must be shared between the connections of the terminal.

30 Accordingly, for a terminal with a plurality of connections with different classes of service, the invention proposes to start by progressively freeing up the high-

performance connections. If this is not sufficient, the equivalent bandwidth of the connections that are not in real time is progressively reduced, and if the equivalent bandwidth of a connection falls below r_{min} , the connection is cleared down. If this is not sufficient, the equivalent bandwidth of the real time connections is progressively reduced to r_{min} in order to maintain the minimum binary bit rate channel.

On the other hand, for a terminal which has only one high-performance connection, the equivalent bandwidth is already established at r_{min} and in this case the terminal is disconnected only if the global bandwidth \times power product exceeds α_{lim}^T .

Also, during the second step, it may happen that the equivalent bandwidth reduction for a small number of terminals is sufficient to return the global bandwidth \times power product to α_{lim}^T . The second step then begins by selecting a terminal to be processed at random.

Capacity increase

A reduction in capacity can be considered as a transient event. Accordingly, at the end of the fading or absorption event, the capacity increases and the system returns to a state of availability of services. The equivalent bandwidth, which was reduced during the capacity reduction, is returned to its normal value.

The first step individually verifies the current value of $\alpha^{(i)}$ against $\alpha_{lim}^{(i)}$. If the current value of $\alpha^{(i)}$ falls below the limit for the terminal, the associated equivalent bandwidth is returned to the nominal equivalent bandwidth established by the CAC function.

The second step globally verifies the current value of α^T against α_{lim}^T . If the current value of α^T falls below the limit for the carrier, the equivalent bandwidth of the terminals is progressively returned to the nominal value established by the CAC function.

The processing to reduce and increase the capacity can be performed automatically by a suitable program that manages modification of the equivalent bandwidth of the terminals and can have the following algorithmic structure, for example.

Individual processing

1) Calculate $\alpha^{(i)}$ from:

$$\alpha^{(i)} = r_i^* \cdot p_i$$

where r_i^* is the nominal equivalent bandwidth established by the CAC function.

$r_i^* = r_i$ when the power margins are not used up, but $r_i^* > r_i$ otherwise, so that r_i^* must be

memorized so that r_i can be returned to its original value after the reduction in capacity.

2) If $\alpha^{(i)} > \alpha_{lim}^{(i)}$, the situation is one of access to the availability of the system. A

reduced equivalent bandwidth r_i is calculated from:

$$r_i = r_i^* \cdot \frac{\alpha_{lim}^{(i)}}{\alpha^{(i)}}$$

5 If $r_i < r_{min}$, the terminal is disconnected from the access network.

3) If $\alpha^{(i)} \leq \alpha_{lim}^{(i)}$, r_i is returned to its nominal value r_i^* .

Collective processing

1) Calculate α_1^T and α_2^T (on each clock cycle) from:

$$\alpha_1^T = \sum_{SKT} r_i^* \cdot p_i \text{ and } \alpha_2^T = \sum_{SKT} r_i \cdot p_i$$

10 2) If $\alpha_2^T < \alpha_{lim}^T$, following the new values of r_i defined during individual processing and the new values of p_i , then the bandwidths r_i are not modified.

3) If $\alpha_2^T \geq \alpha_{lim}^T$, a new value is calculated for the bandwidths r_i for all the terminals t_i , such that:

$$r_i' = r_i \times \frac{\alpha_{lim}^T}{\alpha_2^T}.$$

15 A terminal t_i is disconnected if $r_i' < r_{min}$.

Each time a new value of r_i' is determined, a new value of α_2^T is calculated and the determination of a number r_i' is continued for as long as the condition $\alpha_2^T < \alpha_{lim}^T$ is not verified. The advantage of this solution is that the processing time can be limited as a function of the total number of terminals to be processed (real time processing).

20 Alternatively, each time a terminal t_i is disconnected, α_2^T can be recalculated taking account only of that disconnection (the values r_i are used again and the values r_i' calculated up to this point are ignored), and the whole of the processing restarted.

4) If $\alpha_1^T > \alpha_{lim}^T$, the situation is one of access to the availability of the system. A terminal is chosen for which a reduced equivalent bandwidth r_i is calculated from the

25 equation:

$$r_i = r_i^* \frac{\alpha_{lim}^T}{\alpha^T}$$

The reduction is applied successively for the terminals, but after each calculation of a reduced equivalent bandwidth, α_1^T must be calculated again from the equation:

$$\alpha_1^T = \sum_{SKT} r_i \cdot p_i.$$

If α_1^T becomes less than α_{lim}^T , the reduction of the equivalent bandwidth is halted. If $r_i = r_{min}$ for each value of i , a random choice is made to select the terminals to be disconnected.

- 5) If $\alpha_2^T \leq \alpha_{lim}^T$ and $\alpha_1^T > \alpha_{lim}^T$, r_i is returned to its nominal value r_i^* successively for the terminals which have a reduced equivalent bandwidth. After each equivalent bandwidth is returned to its nominal value, α^T must be calculated again from the equation:

$$\alpha^T = \sum_{SKT} r_i \cdot p_i.$$

- If α^T becomes greater than α_{lim}^T , returning the equivalent bandwidth to its nominal value is halted.

- 6) If $\alpha_1^T \leq \alpha_{lim}^T$, all the values of r_i are returned to their nominal value r_i^* .

II. Terminal to gateway link

The method can be the same as that previously described for the gateway to terminal link, based on conserving the bandwidth \times power product.

- However, because of multiple access problems, the terminal may require a traffic supervisor for redistributing the radio resources allocated by the gateway to transmit traffic data on the gateway to terminal link. A particular logical signaling interface is constructed between the gateway and the terminal to make the reduction of the equivalent bandwidth efficient and to inform the algorithm of the traffic supervisor of the terminal of the reduction of r_i .

The return strategy is generally imposed by the constraint of having a very low capacity cost for transmitting the message informing the terminal of the equivalent bandwidth reduction. The number of reduction or increase steps will generally be reduced.

- The two types of link (uplink and downlink) can be managed separately or in a more or less interleaved manner.

Accordingly, in a first embodiment of the invention, uplink radio resources management is independent of downlink radio resources management, except for situations of disconnection of terminals t_i , the consequences of which are taken into account in both management processes.

5 In a second embodiment, the method can manage uplink radio resources first and then downlink radio resources second, or vice versa, taking into account disconnections resulting from the management process performed first, and then repeating the management process performed first allowing for any disconnections occurring during the management process performed second.

10 Finally, in a third embodiment, uplink radio resources management can be correlated with downlink radio resources management, significantly reducing the bandwidths r_i of a given terminal t_i in the two transmission directions in the same manner.

Of course, the invention is not limited to the embodiment described, which can be modified, in particular from the point of view of the nature of its various components
15 or by substituting technical equivalents, without departing from the scope of the protection afforded to the invention.